Vehicle Value Estimator Using Convolutional Neural Networks

Author(s) : Shafe Assaf, Phalguna Srinivasa Yelandur
Department of Electrical and Computer Engineering, Georgia Institute of Technology

2

Introduction

In the past decade, various resellers have advanced heavily into online sales, with websites such as Amazon, eBay, and Alibaba heavily overtaking traditional retailers. This trend has held true in the used vehicle market as well.
These websites provide comprehensive images of the vehicle for their customers and by viewing images of the vehicle, the user can make decisions as to whether the listed price is reasonable, and make the effort to contact the seller and potentially purchase the vehicle. Since this decision is based on a visual correlation between the vehicle's visual condition and price, this can be automated through neural networks.
Traditional neural networks can be challenging to implement here. Due to the large number of pixel values, traditional neural networks require immense levels of processing power and many layers to create trends to improve performance. Therefore, convolutional neural networks (CNN's) are the best choice to implement this system.

Dataset

The main requirement for the implementation of the neural network is the data set. The dataset contains of 4000 images which has been split into training and testing sets. The images are of white cars obtained from various websites selling pre-owned cars as well as auction sites. All the images are of a single orientation, the left view of the car to maximize efficiency of the neural network and reducing the number of unnecessary variables due to the scale of the dataset. Images obtained from the websites include both the left and right view but have been flipped before being fed into the neural network.
The images are of the size 640X480 initially with the labels of the images being the price of the car obtained from the websites selling them. The dataset contains the images of 5 cars namely Honda Accord, Ford Fusion, Hyundai Sonata, Toyota Camry and Toyota Corolla which were the most popular cars on sale.

Fig. 1. Toyota Camry

Experiments

The network developed contained k-fold testing and validation along with the training layer to improve it's quality.
The training section consists of 4 layers of 32 filters, 2 pooling layers to down sample the convolved images, and some flattening layers at the end to compress the image into a single variable for comparison with the vehicle values. The model is trained using a stochastic gradient descent with a learning rate of 0.01. The best performing loss function was the mean squared logarithmic error, as this allowed the function to start at abnormally high values during problematic k-fold dataset selections. Other loss functions, such as mean squared error and mean absolute error, led to ludicrously high values and nan errors.
The evaluation portion for the validation data sets simply scores the model. The evaluation for the test data does multiple steps. First it evaluates the model and then it creates label estimates based off of the data set, which are the vehicle's estimated values. The data set was normalized from 0 to 1, and then multiplied with by the current maximum value of the vehicle minus its lowest value and then added to the minimum offset.
The implementation of the network with no pre processing before being fed into yielded better results when compared to the any pre processing filters applied on the training images. The graph of estimated value vs actual value is as shown in figure 4.
The high variation in the estimated values of the cars can be attributed to the lack of visual factors affecting the price of the car.
The use of a color filter to remove background information resulted in the loss of vital information about the damage to the car (as shown in figure 3). This affected the estimated values of the car by a huge margin. The plot for the estimated values obtained using pre processing techniques are in figure 5.
The results obtained using k-fold testing (5 fold and 7 fold) are shown in figure 6.

Figure 2

Figure 3

Images with color filter applied to remove background information

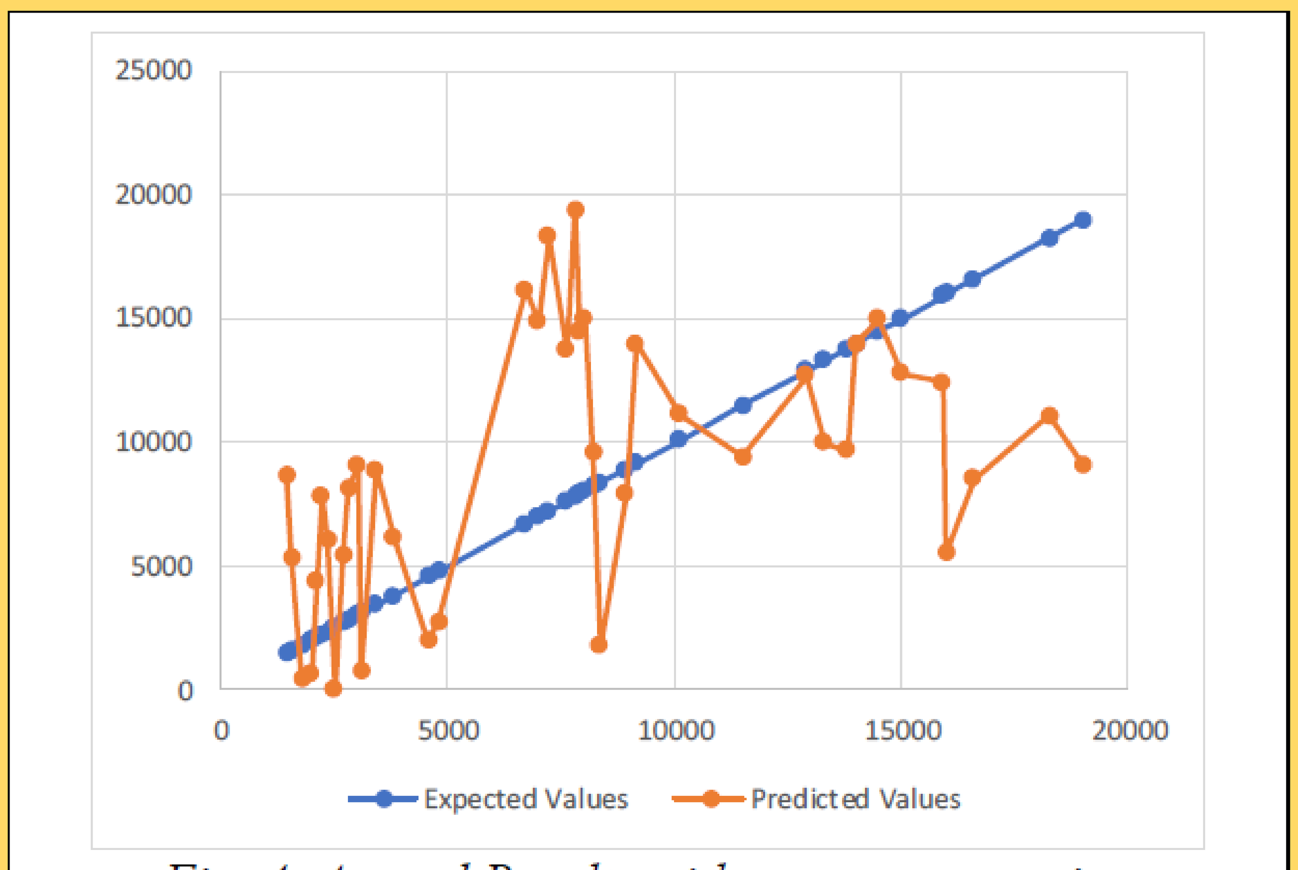

Fig. 4. Accord Results with no pre-processing

Results with no pre-processing done for Honda Accord

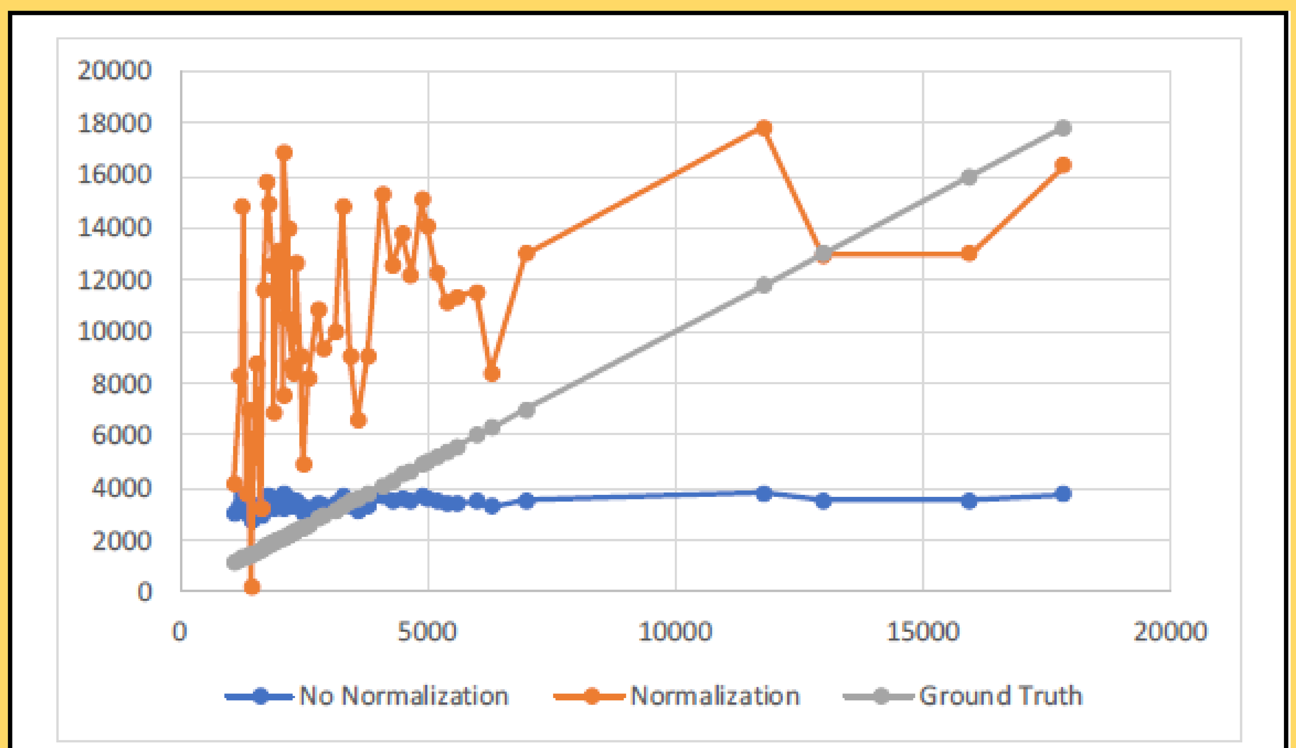

Fig. 5. Camry results with pre-processing

Results with color filter applied to images of Toyota Camry

[0.507, 0.640, 0.794, 75.636, 0.634]
a) 5-fold results

[0.674, 0.686, 74.340, 0.737, 0.636, 0.781, 0.712]
b) 7-fold results

Fig. 6. K-fold results for color filter

Conclusion

The results showed that the unfiltered images had strong results for the lower priced and more heavily damaged vehicles, but had high fluctuations for higher priced vehicles. This can be due to the factor of heavy physical damage sustained by the vehicles resulting in their prices being of near scrap metal. values, which can be easily detected using the image. With the increase in the price ranges, the physical damage sustained reduces and hence the detection of defects proves to be more difficult. This result can be improved with the increase in the amount of data fed into the network.
Meanwhile, color filtering seemed to eliminate the outliers within higher priced vehicles, but inadvertently eliminated the subtleties of color within damaged portions of the vehicle, leading to poor low light performance.
In conclusion, The network performs well on greyscale images with cars containing high physical damage.